Exploratory key nodes in the inventor-author knowledge diffusion network

Guijie Zhang*
School of Management Science and Engineering, Shandong University of Finance and Economics, Jinan 250014, China
Corresponding author
zgjzxmtx@163.com

Fangfang Wei*
Business School, University of Jinan, Jinan 250002, China
Corresponding author
weifftju@163.com

Zhen Shao
School of Management, Harbin Institute of Technology, Harbin 150001, China
shaozhenlily@gmail.com

Abstract

This paper aims to mine the key nodes in the process of knowledge flow from literatures of science and technology journals to technology patents on the community level. Based on the citation of technological patents to literatures of scientific journals and the cooperation among the researchers, this paper builds the knowledge flow network from the angle of spatial dimension. Then employing the extensity centrality-Newman and the commonly used degree indexes, this paper excavates and analyses the nodes which occupy important positions among communities in the knowledge flow network. After that, this paper puts forward suggestions on how to make full use of the key nodes' role of bridge to promote knowledge flow from literatures of science and technology journals to technology patents.

1. Introduction

Knowledge is the core competitive advantage of enterprises and can be used to achieve commercial benefits in a constantly changing global economic environment[1]. As one of the most crucial sources of value creation, knowledge flow plays an irreplaceable role and accordingly, it has received widespread attention[2-4]. Through effective knowledge flow mechanisms, large quantities of high-quality knowledge will realize their value[5].

There have been many efforts to characterize the nature of knowledge flow from multiple perspectives of detail, namely: among scientists[6, 7], among patents[3, 8], among patents and journals[9-11], among journals[12-14], among disciplines[15, 16], among institutions[17], and among countries[18-20], etc. The prior researches provide a comprehensive understanding of the nature of knowledge flow; in particular, enrich the potential utility of analyzing its processes. Due to the comprehensive coverage of information, patent citation is one of the most popular measures of knowledge flows in the information science and technology[6, 21], management science[2, 19], and research policy literatures[3, 22], etc.

Social network can establish connections among researchers and plays a facilitative part in enhancing knowledge flow[14, 23]. Social network analysis is a powerful means to investigate the interactions and patterns of knowledge flows among group members[20]. Identifying the significance of nodes in the network is a classical topic in the field of informetrics[7, 24, 25]. Plenty of indictors are proposed from diverse perspectives to analysis the properties of nodes in the network, such as the directed h-degree[26], C-index[27], success-index[28], Cext-N[7], I-index[25], D-index[24], Basic Research Competitiveness Index[29], et al. As every researcher in the network has his unique knowledge accumulation, a wealth of knowledge flow among them could promote the progress of scientific research and reinforce research quality[30].

Community structure is one of the most important properties of social network and in particular the knowledge flow networks[7, 31]: the nodes in networks often cluster into tightly-knit groups with a high density of within-group edges and a lower density of between-group edges[32]. Different communities have diverse characteristics which can complement each other. Because of its significance in social network analysis, lots of scholars in a variety of fields have paid attention to the identification of community structure and numerous classic methods have been developed to obtain optimal solutions, such as Girvan-Newman algorithm[32], VOS Clustering[33], topic oriented community detection approach[34], information-theoretic approach for detecting communities[35], degree-corrected block model[36], integrating center locating and membership optimization algorithm[37], and improved algorithm based on the random graph models[38], et al.

URI: http://hdl.handle.net/10125/50422
ISBN: 978-0-9981331-1-9

The status of individuals in the scientific community not only has a significant effect on his/her reputation and performance but also influence knowledge flow of the network[39, 40]. Based on the community detection algorithms, some scholars take efforts to establish indicators to examine the nodes' role on knowledge flow. Among these indicators, the index named extensity centrality can measure the distribution of knowledge flow among different communities, and analyze whether the researchers collaborate with others from multiple communities[31]. The idea of a researcher's cooperation extensity perspective contributes to a better understanding of scholarly structures and activities, and can expand co-operation activity measurement of researchers. In computing the strength of collaborative ties among authors in co-authorship network, which is closely related to the extensity centrality, Lv and Feng (2009) chose Salton's measure which employs the information of how many papers each pair of scientists has collaborated[31]. Considering the information used in the measure, Zhang et al. (2014) proposed the extensity centrality-Newman which not only uses the information of how many papers each pair of scientists has collaborated, but also considers how many coauthors they have on each paper[7]. Taking the co-authorship network in the field of management information system as an example, the results indicate that the Newman's measure is more suitable than Salton's measure in computing the extensity centrality[7].

The previous efforts mining key nodes on the community level have mostly focused on the co-authorship network while few attentions have been paid to the inventor-author knowledge flow network[7, 31]. Because of the great value of knowledge flow between literatures and patents, such as measuring science-technology interactions[41], promoting technological innovation[42], calculating science linkage[10], and analyzing the time delay between scientific research and technology patents[11], etc, finding the key nodes which serve an important function in the knowledge flow process among different communities has great significance. The objective of this paper is to fill this gap by employing the extensity centrality-Newman to detect the important nodes in the inventor-author knowledge flow network, to better understand these nodes' function in the knowledge flow process as well as to provide suggestions to government for their policy decisions.

The reminder of this paper is organized as follows. In 'Method' section, we introduce the extensity centrality-Newman model. This is followed by 'Data collection' section, which describes the data collection procedure and provides brief descriptive statistics. Afterwards, the 'Results' section conducts two case studies and discusses the results. In 'Conclusion and discussions' section, we illustrate the conclusion and provide some suggestions.

2. Method

In this paper, the Lambda sets put forward by Borgatti et al. (1990)[43] will be chosen for community detection, since this measure can perfectly depict the cohesiveness and stability of communities[31]. In the calculation of the strength of collaborations between scholars, Salton measure and Newman measure are the most well-known indicators[7]. According to Salton's measure, the more cooperation among scholars, the closer the relationship among them is. The collaboration strength between two scholars is described as follows:

$$r_{ij} = h_{ij}/\sqrt{h_i * h_j} \quad (1)$$

where hij is the number of papers coauthored by scholar i and j; hi and hj are numbers of their separate papers.

In addition, collaborative ties weight inversely according to the number of co-researchers of each paper. To account for this factor, Newman (2001a, b) proposes a new indicator to measure the strength of the collaboration between scholar i and j[44, 45].

$$w_{ij} = \sum_k \frac{\delta_i^k \delta_j^k}{n_k - 1} \quad (2)$$

where δ_i^k is 1 if scholar i is a co-researcher of patent/paper k and zero otherwise, and n_k is the number of co-researcher of patent/paper k. The Newman measure is more reasonable than Salton measure as it includes more information during the measurement. Consequently, in our calculation of the strength of collaborative ties between scholars, we will apply Newman's measure.

Inspired by the entropy centrality[46, 47], scholars proposed the idea of researcher's extensity. Assuming that a knowledge flow network has n researchers and c communities, the sum of a researcher's strength of knowledge flow ties with others in community k is

$$\mathrm{sum}_k(i) = \sum_j w_{ij} \quad (3)$$

where j is one of researcher i's collaborators in community k. The proportion of researcher i's collaboration with others in community k is

$$P_k(i) = \frac{\mathrm{sum}_k(i)}{\sum_{j=1}^{n} w_{ij}} \quad (4)$$

Based on Shannon's measure of information, Tutzauer (2007)[47] proposed the extensity centrality of researcher i as follows:

$$C_{ext}(i) = -\sum_{h=1}^{c} P_h(i) \log P_h(i) \quad (5)$$

Formula 5 does not contain the universal case that an author does not belong to any community. Lv and

Feng (2009)[31] regard those collaborators as a third-party group and make the improvement in formula 4 as follows:

$$C_{ext}(i) = -\sum_{h=1}^{c} P_h(i)\log P_h(i) - (1 - \sum_{h=1}^{c} P_h(i))\log(1 - \sum_{h=1}^{c} P_h(i)) \quad (6)$$

According to formula 6, we can compute researchers' extensity centrality, and the logarithmic base sets 2 in the present study.

3. Data collection

The dependencies of scientific research and technology patents vary in the diverse technological fields[48]. Among these fields, pharmaceuticals and biosensor are closely related to scientific achievements. Consequently, scientific literature plays an important role in their practical development[48, 49]. There are plenty of patents in these fields, which is essential for researchers to draw a clear picture of the knowledge flow between scientific research and technology patents; therefore, they are ideal fields for this investigation[50].

All the patent information of this paper comes from the USPTO website (http://patft.uspto.gov/netahtml/PTO/index.html), which owns rich patent information and has been used by massive previous studies[9, 10, 51]. We precede the data collection process employing a modular keyword searching strategy[10, 52, 53]. The search commands used in this paper are as follows: ((TTL/"Pharmaceuticals" OR ABST/"Pharmaceuticals") AND ISD/20110101->20131231); ((TTL/"biosensor" OR ABST/"biosensor") AND ISD/20110101->20131231). The numbers of patent in the field of pharmaceuticals and biosensor are 423 and 376, respectively. To ensure the follow-up study smoothly, we conducted a thorough cleaning process by deleting the extra information, and only scientific papers retained. After that, the total numbers of cited articles are 8914 and 4378, respectively. Among these articles, the numbers of self-citations by applicant (articles written by applicant) are 358 and 208, respectively. Then we collect the inventors and authors of these patents and citations, unify their names format, and build the inventor-author knowledge flow network.

4. Results

Employing the Pajek program, we provide the visualization of network of the field of pharmaceutical which scale is 840, as is shown in Figure 1. After analyzing, 58 components are found in this network and the two largest components are constituted of 107 and 81 points, respectively, occupying 22.38 % of the whole. In addition, there are still many other small components which scales are all less than 50. In order to learn the detail of the number of components, we present table 1 which is shown as follows.

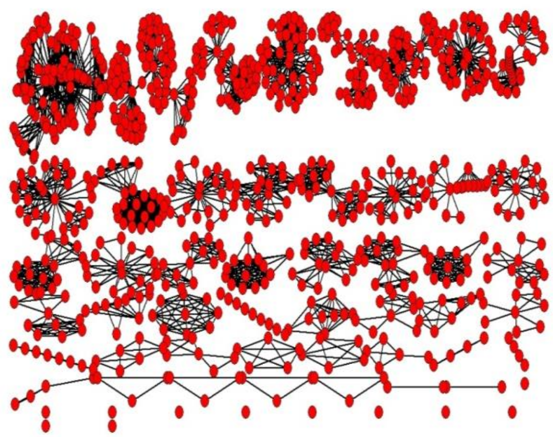

Figure 1 Visualization of the network structure of pharmaceuticals

Table 1 Detail of components in Pharmaceuticals

Interval	[50,)	[40,50)	[30,40)	[20,30)	[10,20)	[1,10)
Scale and proportion	188(22.38%)	130 (15.48%)	133 (15.83%)	46 (5.48%)	209 (24.88%)	134 (25.95%)

Then using Pajek program, we analysis the network composed by inventors and authors in the field of biosensor, as shown as Figure 2. After component analysis, we find 180 components in total. The scales of the two largest components are 69 and 58, respectively, which occupy 15.96% of the whole network. As for this network, there also exit many relatively small components with the scales all less than 40, occupying 84.04% totally. In order to learn the distribution of the components in the field of biosensor, we present the table 2 which is shown as follows.

Table 2 Detail of components in biosensor

Interval	[50,)	[40,50)	[30,40)	[20,30)	[10,20）	[1,10)
Scale and proportion	127(15.96%)	0 (0%)	36(7.75%)	134(18.90%)	169(19.25%)	334(38.15%)

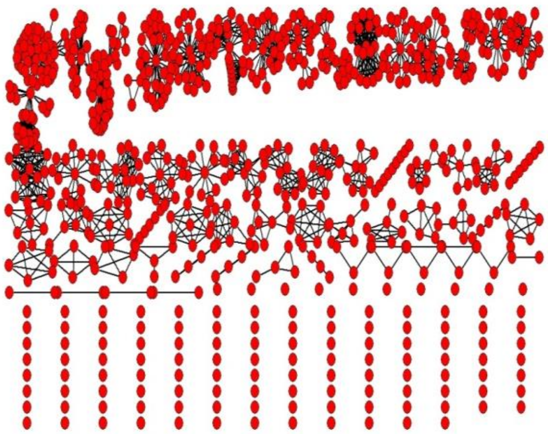

Figure 2 Visualization of the network structure of biosensor

From Fig. 1 and Fig. 2, and table 1 and table 2, we can find out that although the scales of the field of pharmaceutical and biosensor are relatively closer, the numbers of components in these two areas are quite different: in the field of pharmaceutical the components' number is 58 while that of biosensor is 180. The scale of the biggest component in the field of pharmaceutical is 107 which is much larger than that of biosensor which is 69. Relatively speaking, the inventors and authors in the field of pharmaceutical have more connections than that of biosensor. Generally, the social network usually consists of a great many components, and giant component is a universal phenomenon in an academic network. Due to the characteristics of the small network are not obvious and the giant component may signify the core of mainstream research activity [7, 54], the common procedure is to narrow the analysis to the biggest one[55]. Thus in this paper, we only pay attention to the largest components in these two fields, with the scales of 107 and 69, respectively, which are large enough for our further analysis.

In order to learn the distribution of the largest component in the field of pharmaceutical, using the Pajek and Ucinet program, we illustrate the details of network graph of this component, as is shown in Fig. 3. Then employing the Pajek and Java program, we observe the value of the extensity centrality-Newman (Cext-N), also analyze the classical indices of this component, such as degree centrality (CD), betweenness centrality (CB), closeness centrality (CC) and eigenvector centrality (CE). To visually observe the differences in various indices, we arrange the first 20 researchers as the representative with their corresponding number according to the descending order of each index. The results are shown in table 3. In the research process, the Lambda value can be selected according to the actual situation. For example, when the value of lambda is chosen 1, the network composed by the researchers will form one lambda set. If the Lambda value is too small, the degree of differentiation is not high enough; in contrast, if the Lambda value is too large, the network will be too dispersed to analysis. As the largest component of the field of pharmaceutical contains 107 nodes, which is relatively large, here the lambda is chosen 6.

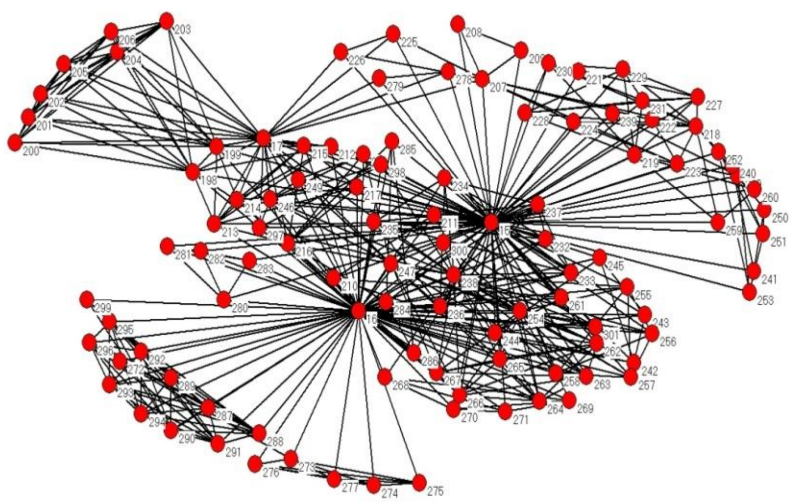

Figure 3 Visualization of the network structure of component one of pharmaceuticals (107 nodes)

Table 3 Index values of the first 20 authors of component 1 in pharmaceuticals

Cext-N	C_D	C_C	C_B	C_E
1.000 (207)	82 (15)	81.538 (15)	53.859 (15)	0.420 (15)
1.000 (224)	71 (16)	75.177 (16)	40.418 (16)	0.406 (16)
0.957 (300)	27 (17)	57.297 (17)	7.322 (17)	0.223 (238)
0.957 (301)	26 (238)	56.989 (238)	3.289 (198)	0.182 (17)
0.947 (261)	16 (254)	53.807 (247)	3.289 (199)	0.164 (247)
0.901 (222)	15 (247)	53.535 (198)	1.495 (238)	0.159 (254)
0.883 (223)	14 (216)	53.535 (199)	0.434 (207)	0.131 (216)
0.865 (15)	14 (198)	53.535 (216)	0.344 (218)	0.127 (262)
0.846 (216)	14 (199)	52.475 (214)	0.326 (216)	0.127 (263)
0.801 (210)	13 (218)	52.475 (246)	0.31 (247)	0.127 (264)
0.801 (211)	11 (287)	52.475 (213)	0.225 (254)	0.127 (265)
0.796 (16)	11 (289)	52.217 (254)	0.126 (220)	0.127 (266)
0.777 (220)	11 (290)	51.707 (249)	0.103 (213)	0.127 (267)
0.766 (242)	11 (292)	51.707 (248)	0.103 (214)	0.120 (246)
0.766 (243)	11 (295)	51.707 (212)	0.102 (219)	0.108 (198)
0.766 (245)	10 (207)	51.456 (297)	0.054 (244)	0.108 (199)
0.697 (208)	10 (213)	51.456 (298)	0.045 (246)	0.108 (213)
0.697 (209)	10 (214)	50.718 (264)	0.036 (273)	0.108 (214)
0.657 (284)	10 (219)	50.718 (262)	0.015 (289)	0.097 (232)
0.656 (285)	10 (246)	50.718 (267)	0.015 (295)	0.097 (233)

Table 3 indicates that the ranking results according to different indices of scholars are also different. The results in table 3 suggest that the extensity centrality-Newman is quite different from that of the classical indices. After analyzing, we can find that the nodes numbered 15, 207, 216 appear in almost every indexes, which means no matter from which angle, these researchers are very important in the network. Employing the SPSS program, we calculate correlation of these five indices to comprehensively examine the relationship between them. The correlation coefficient between the index of the extensity centrality-Newman and betweenness centrality is 0.239*, and that between the index of the extensity centrality-Newman and closeness centrality is 0.233*. There is no significant correlation between the extensity centrality-Newman and degree centrality, and between the extensity centrality-Newman and eigenvector centrality. Additionally, the classical indices such as degree centrality, betweenness centrality, closeness centrality and eigenvector centrality show significant correlation with each other at the 0.01 level, and the correlation coefficient are all above 0.75**, indicating the close relationship between these indices.

To learn the distribution of nodes in the largest component, we present Fig.3, which clearly illustrates that the nodes such as 15, 16, and 17 have a lot of connections with others, occupying the central position of the network. The values of the degree centrality of these nodes are very high, indicating that these nodes are very active in the network. Take the node numbered 15 as an example. Node 15 corresponds to Professor Peter Palese, who is the Chair of the Department of Microbiology at the Icahn School of Medicine and an authoritative expert in the field of RNA viruses. Professor Peter Palese is not only a famous scholar, but also a good inventor, who has published more than 400 scientific periodical literatures and holds a number of patents. The value of extensity centrality-Newman of this node is 0.865, which means that Professor Peter Palese plays an important role in knowledge communication in the field of pharmaceuticals. Besides, he serves in several National Academies of Science, mastering a large number of resources. Accordingly, he communicates with different subgroups in the network, and plays a bridge role in the process of knowledge flow.

Employing the Pajek and Ucinet program, we illustrate the network diagram of the largest component in the field of biosensor, as is shown in Fig. 4. Then using the Pajek and Java program, we observe the value of the extensity centrality-Newman (Cext-N), and analyze the classical indices of component 69, such as degree centrality (C_D), betweenness centrality (C_B), closeness centrality (C_C) and eigenvector centrality (C_E). To clearly observe the differences of various indices, we arrange the first 10 inventors and authors with their corresponding number according to the descending order of each index. The results are shown in Table 4. Here, the lambda is chosen 4.

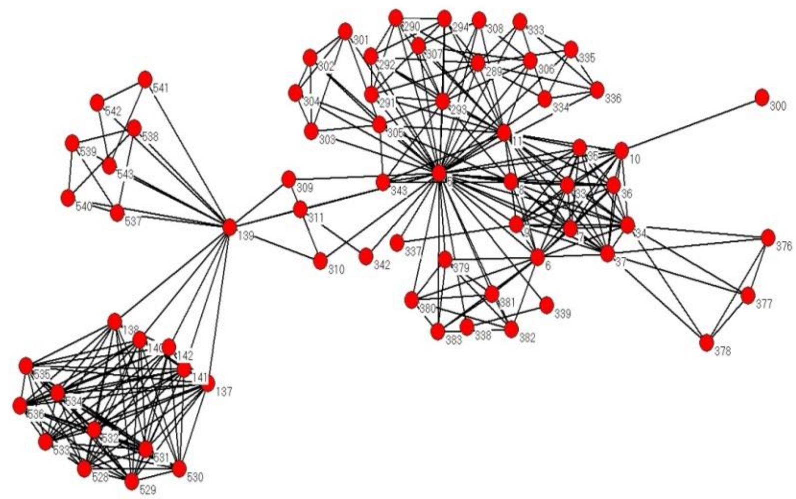

Figure 4 Visualization of the network structure of component one of biosensor (69 nodes)

Table 4 Index values of the first 10 authors of component 1 in biosensor

Cext-N	C_D	C_C	C_B	C_E
0.586 (311)	43 (5)	66.667 (5)	71.071 (5)	0.270 (137)
0.560 (139)	22 (11)	54.839 (139)	48.156 (139)	0.270 (138)
0.512 (9)	16(139)	47.222 (311)	4.346 (142)	0.270 (140)
0.383 (10)	16 (6)	47.222 (11)	4.346 (140)	0.270 (141)
0.267 (5)	14 (8)	46.897 (310)	4.346 (137)	0.270 (142)
0.229 (541)	14 (34)	46.897 (309)	4.346 (138)	0.261 (528)
0.229 (542)	14 (37)	45.333 (6)	4.346 (141)	0.261 (529)
0.229 (543)	14(137)	44.737 (8)	4.214 (37)	0.261 (535)
0.111 (338)	14(138)	44.737 (37)	4.214 (34)	0.261 (536)
0.111 (339)	14(140)	44.737 (34)	3.424 (11)	0.261 (532)

Table 4 indicates that the ranking results according to different indices of scholars are also different. These results suggest that the extensity centrality-Newman is quite different from that of the classical indices. After analyzing, we can find that the nodes numbered 5, 139, 311 appear in almost every indexes, which means no matter from which point of view, these researchers are very important in the network. Employing the SPSS program, we calculate correlation of these five indices to comprehensively examine the relationship between them. The correlation coefficient between the index of the extensity centrality-Newman and closeness centrality is 0.438**, showing significant correlation with each other at the 0.01 level. There is no significant correlation between the extensity centrality-Newman and degree centrality, betweenness centrality, and eigenvector centrality. Additionally, the correlations of the classical indices are as follows: degree centrality and closeness centrality is 0.342**; degree centrality and betweenness centrality is 0.688**; degree centrality and eigenvector centrality is 0.439**; closeness centrality and betweenness centrality is 0.466**; closeness centrality and eigenvector centrality is 0.432**, indicating the close relationship between these indices. In addition, there is no significant correlation between eigenvector centrality and betweenness centrality.

Fig.4 clearly illustrates that the node 311 (D. Lansing Taylor) links the two sets of the left (including 22 nodes) and right sides (including 45 nodes) of the network. The value of the extensity centrality-Newman of this node is very high and ranks first in the network, occupying the central position of the network. The node 139 (Alan S. Waggoner) communicates with different subgroups in the network which are composed by (537, 538, 539, 540, 541, 542, 543) and (137, 138, 140, 141, 142, 528, 529, 530, 531, 532, 533, 534, 535, 536), respectively. This indicates that Alan S. Waggoner is very active in the network, making the center of the value of this node is very high (0.560), ranked in the second in the network. And thus this researcher plays a bridge role in the process of knowledge flow.

Generally speaking, among the classic indices, there is a high correlation between the extensity centrality-Newman and closeness centrality, and no general correlation between the extensity centrality-Newman and the other centralities, including degree centrality, betweenness centrality, and eigenvector centrality. This can be explained from the definition of the extensity centrality-Newman and closeness centrality. The closeness centrality is proposed to measure the importance of the node in the network from the perspective of communication, while the extensity centrality-Newman concerns the degree of closeness contact with different subgroups in the network. There is a certain degree of similarity between these two indexes. Accordingly, from the above results we can find that the nodes with high extensity centrality-Newman value tend to have high values of the closeness centrality. Not vice versa. This is because the nodes with high value of betweenness centrality maybe at the core of the network and some shortest paths through it, but this node does not necessarily connect different subgroups, makes the values of the extensity centrality-Newman of these nodes not high enough.

5. Conclusion and discussion

In this paper, based on the citation relationship between technological patents and scientific literatures and the cooperative relationship among researchers, we constructed the inventor-author network and quantitatively investigated the key nodes of the knowledge flow network on community level. Our study makes both empirical and theoretical contributions to the current understanding of key nodes identification. In terms of theoretical contributions, this paper effectively expands the application scope of the extensity centrality-Newman, which focuses on measuring whether a researcher collaborates with others in and out of the community, from co-author network to inventor-author network. In addition, this paper further discusses the relationship between the extensity centrality-Newman and the classical indexes, such as degree centrality, betweenness centrality, closeness centrality and eigenvector centrality.

In terms of empirical contribution, mining key nodes in knowledge flow network on community level have important values in scientific research. Firstly, as the key inventors/authors in the network have a good professional knowledge reserve in the field, the patent office can employ them as consultants to participate in the process of patent examination. With the help of these researchers, the patent office can confirm the actual value of the patent application and the rationality of their rights in less time, thus to speed up the process of patent examination. Secondly, the excavation of the key nodes can help the researchers more targeted in finding collaborators, which is meaningful for accelerating the knowledge flow process and promote the invention of technology. Thirdly, the extensity centrality-Newman can be used as one of the indexes for evaluating the comprehensive ability of researchers. The extensity centrality-Newman serves as an effective complement to existing indexes, and the researcher with high extensity centrality-Newman can effectively integrate the resources of different researchers, promote the knowledge flow, improve resource utilization, and increase researchers' output.

6. References

[1] Lai, C.H., Applying knowledge flow mining to group recommendation methods for task‐based groups. Journal of the Association for Information Science & Technology, 2015. 66(3): p. 545‐563.

[2] Roach, M. and W.M. Cohen, Lens or Prism? Patent Citations as a Measure of Knowledge Flows fromPublic Research. Management Science, 2013. 59(2): p. 504.

[3] Roper, S. and N. Hewitt-Dundas, Knowledge stocks, knowledge flows and innovation: Evidence from matched patents and innovation panel data. Research Policy, 2015. 44(7): p. 1327-1340.

[4] Azagra-Caro, J.M. and D. Consoli, Knowledge flows, the influence of national R&D structure and the moderating role of public–private cooperation. Journal of Technology Transfer, 2016. 41(1): p. 152-172.

[5] Yu, G., M.Y. Wang, and D.R. Yu, Characterizing knowledge diffusion of Nanoscience & Nanotechnology by citation analysis. Scientometrics, 2010. 84(1): p. 81-97.

[6] Yan, E., Y. Ding, and X. Kong, Monitoring knowledge flow through scholarly networks. Proceedings of the American Society for Information Science & Technology, 2013. 49(1): p. 1-5.

[7] Zhang, G., et al., Cext-N index: a network node centrality measure for collaborative relationship distribution. Scientometrics, 2014. 101(1): p. 291-307.

[8] Chen, L., Do patent citations indicate knowledge linkage? The evidence from text similarities between patents and their citations. Journal of Informetrics, 2017. 11(1): p. 63-79.

[9] Huang, M.H., et al., Exploring temporal relationships between scientific and technical fronts: a case of biotechnology field. Scientometrics, 2014. 98(2): p. 1085-1100.

[10] Li, R., et al., Patent citation analysis: Calculating science linkage based on citing motivation. Journal of the Association for Information Science & Technology, 2014. 65(5): p. 1007-1017.

[11] Zhang, G., et al., Analyzing the time delay between scientific research and technology patents based on the citation distribution model. Scientometrics, 2017. 111: p. 1-20.

[12] Zhou, P. and L. Leydesdorff, A comparison between the China Scientific and Technical Papers and Citations Database and the Science Citation Index in terms of journal hierarchies and interjournal citation relations. Journal of the Association for Information Science & Technology, 2007. 58(2): p. 223-236.

[13] Zhou, P., X. Su, and L. Leydesdorff, A comparative study on communication structures of Chinese journals in the social sciences. Journal of the Association for Information Science & Technology, 2010. 61(7): p. 1360–1376.

[14] Gao, X. and J. Guan, Network model of knowledge diffusion. Scientometrics, 2012. 90(3): p. 749-762.

[15] Tsay, M.Y., Knowledge flow out of the domain of information science: a bibliometric and citation analysis study. Scientometrics, 2015. 102(1): p. 487-502.

[16] Zhu, Y. and E. Yan, Dynamic subfield analysis of disciplines: an examination of the trading impact and knowledge diffusion patterns of computer science. Scientometrics, 2015. 104(1): p. 335-359.

[17] Liu, X., et al., Modeling knowledge diffusion in scientific innovation networks: an institutional comparison between China and US with illustration for nanotechnology. Scientometrics, 2015. 105(3): p. 1953-1984.

[18] Hassan, S.U. and P. Haddawy, Measuring international knowledge flows and scholarly impact of scientific research. Scientometrics, 2013. 94(1): p. 163-179.

[19] †, Z.C. and J.G. †, The core-peripheral structure of international knowledge flows: evidence from patent citation data. R & D Management, 2015. 46(1): p. 62–79.

[20] Darvish, H. and Y. Tonta, Diffusion of nanotechnology knowledge in Turkey and its network structure. Scientometrics, 2016. 107(2): p. 569-592.

[21] Lin, C., J.C. Wu, and D.C. Yen, Exploring barriers to knowledge flow at different knowledge management maturity stages. Information & Management, 2012. 49(1): p. 10-23.

[22] Wu, C.Y. and J.A. Mathews, Knowledge flows in the solar photovoltaic industry: Insights from patenting by Taiwan, Korea, and China. Research Policy, 2012. 41(3): p. 524-540.

[23] Jansen, D., R.V. Görtz, and R. Heidler, Knowledge production and the structure of collaboration networks in two scientific fields. Scientometrics, 2010. 83(1): p. 219-241.

[24] De Sordi, J.O., M.A. Conejero, and M. Meireles, Bibliometric indicators in the context of regional repositories: proposing the D-index. Scientometrics, 2016. 107(1): p. 1-24.

[25] Sahoo, S., Analyzing research performance: proposition of a new complementary index. 2016: Springer-Verlag New York, Inc. 489-504.

[26] Zhao, S.X. and F.Y. Ye, Exploring the directed h - degree in directed weighted networks. Journal of Informetrics, 2012. 6(4): p. 619-630.

[27] Yan, X., L. Zhai, and W. Fan, C-index: A weighted network node centrality measure for collaboration competence. Journal of Informetrics, 2013. 7(1): p. 223-239.

[28] Egghe, L., Impact coverage of the success-index. Journal of Informetrics, 2014. 8(2): p. 384-389.

[29] Zhang, Y., et al., Modelling the Basic Research Competitiveness Index (BR-CI) with an application to the biomass energy field. Scientometrics, 2016. 108(3): p. 1-21.

[30] Liao, C.H., How to improve research quality? Examining the impacts of collaboration intensity and member diversity in collaboration networks. Scientometrics, 2011. 86(3): p. 747-761.

[31] Lu, H. and Y. Feng, A measure of authors' centrality in co-authorship networks based on the distribution of collaborative relationships. Scientometrics, 2009. 81(2): p. 499-511.

[32] Girvan, M. and M.E.J. Newman, Community Structure in Social and Biological Networks. Proceedings of the National Academy of Sciences of the United States of America, 2001. 99(12): p. 7821.

[33] Waltman, L., N.J.V. Eck, and E.C.M. Noyons, A unified approach to mapping and clustering of bibliometric networks. Journal of Informetrics, 2010. 4(4): p. 629-635.

[34] Zhao, Z., et al., Topic oriented community detection through social objects and link analysis in social networks. Knowledge-Based Systems, 2012. 26: p. 164-173.

[35] Li, Y., et al., An entropy-based social network community detecting method and its application to scientometrics. Scientometrics, 2015. 102(1): p. 1003-1017.

[36] Qu, Y., W. Shi, and X. Shi, Inferring overlapping community structure with degree-corrected block model. Physica A Statistical Mechanics & Its Applications, 2015. 419(419): p. 48-54.

[37] Li, H.J., et al., Fast and accurate mining the community structure: integrating center locating and membership optimization. IEEE Transactions on Knowledge & Data Engineering, 2016. 28(9): p. 2349-2362.

[38] Qu, Y., W. Shi, and X. Shi, An improved algorithm for generalized community structure inference in complex networks. Physica A Statistical Mechanics & Its Applications, 2017. 478.

[39] Ductor, L., et al., Social networks and research output. Review of Economics & Statistics, 2011. 96(5): p. 936-948.

[40] Letina, S., Network and actor attribute effects on the performance of researchers in two fields of social science in a small peripheral community. Journal of Informetrics, 2016. 10(2): p. 571-595.

[41] Wang, G. and J. Guan, Measuring science-technology interactions using patent citations and author-inventor links: an exploration analysis from Chinese nanotechnology. Journal of Nanoparticle Research, 2011. 13(12): p. 27–42.

[42] Park, H.-W. and S.-H. Suh. Scientific and technological knowledge flow and technological innovation: quantitative approach using patent citation. in Technology Management in the Energy Smart World. 2011.

[43] Borgatti, S.P., M.G. Everett, and P.R. Shirey, LS sets, lambda sets and other cohesive subsets. Social Networks, 1990. 12(4): p. 337-357.

[44] Newman, M.E., Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2001. 64(2): p. 016132.

[45] Newman, M.E., The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences of the United States of America, 2000. 98(2): p. 404-409.

[46] Rner, K., et al., Studying the emerging global brain: Analyzing and visualizing the impact of co-authorship teams: Research Articles. Complexity, 2005. 10(4): p. 57-67.

[47] Tutzauer, F., Entropy as a measure of centrality in networks characterized by path-transfer flow. Social Networks, 2007. 29(2): p. 249-265.

[48] Verbeek, A., K. Debackere, and M. Luwel, Science cited in patents: A geographic "flow" analysis of bibliographic citation patterns in patents. Scientometrics, 2003. 58(2): p. 241-263.

[49] Zhao, Q. and J. Guan, Modeling the dynamic relation between science and technology in nanotechnology. Scientometrics, 2012. 90(2): p. 561-579.

[50] Finardi, U., Time relations between scientific production and patenting of knowledge: the case of nanotechnologies. Scientometrics, 2011. 89(1): p. 37.

[51] Shin, J.C., S.J. Lee, and Y. Kim, Knowledge-based innovation and collaboration: a triple-helix approach in Saudi Arabia. Scientometrics, 2012. 90(1): p. 311-326.

[52] Guan, J. and N. Liu, Invention profiles and uneven growth in the field of emerging nano-energy. Energy Policy, 2015. 76(1): p. 146-157.

[53] Zhang, G., et al., Improving the publication delay model to characterize the patent granting process. Scientometrics, 2017. 111(3): p. 1-17.

[54] Fatt, C.K., E.A. Ujum, and K. Ratnavelu, The structure of collaboration in the Journal of Finance. Scientometrics, 2010. 85(3): p. 849-860.

[55] Groh, G. and C. Fuchs, Multi-modal social networks for modeling scientific fields. 2011: Springer-Verlag New York, Inc. 569.